Implications of Pension Illiteracy for Labor Supply and Redistribution*

Johan Gustafsson†

June 29, 2021

Abstract

This paper explores the effects of pension illiteracy on aggregate labor supply and the redistributive performance of public pension systems. I consider an overlapping generations model in continuous time populated with individuals who differ in labor productivity and pension literacy. Agents suffering from pension illiteracy fail to fully account for the structure of the pension system when planning their economic behavior over the life cycle. In particular, I assume that myopic agents treat changes to replacement income as exogenous in the active-retired trade-off and contributions to the pension system as a pure labor income tax. I find that pension illiteracy can negatively impact aggregate labor supply and increase earnings inequality and lifetime income inequality. This suggests that pension illiteracy may limit the efficiency gains of increasing the correlation between individual contributions and benefits, making the equity-efficiency trade-off difficult to characterize in the context of pension reforms.

Keywords: Labor supply, Myopia, Public pension

JEL Codes: H55, J22, J26

*Work in progress. Please do not cite nor redistribute.

†E-mail: johan.a.gustafsson@umu.se

1 Introduction

Contemporary public pension systems provide several means of redistributing income over time, between states of the world and individuals. These transfer mechanisms add complexity to economic trade-offs by introducing new implicit and explicit (dis-)incentives for labor supply and savings decisions. Individuals may subsequently find it difficult to fully comprehend how their economic behavior interacts with the public pension system.

Indeed, the Swedish Pensions Agency claims that a common misconception among individuals is that their actions have no effect on their future retirement benefits (Pensionsmyndigheten, 2020). This is remarkable since the Swedish pension system is largely earnings-based, with a highly predictable link between individual contributions and entitled pension income following the introduction of notional pension accounts in the 1990s.[1] Such manifestation of financial illiteracy, or more specifically *"pension illiteracy"*, risks undermining the effectiveness of public pension policy (Bucher-Koenen et al., 2019).[2]

This paper studies the implications of pension illiteracy for labor supply behavior and the redistributive performance of public pension systems. In particular, the purpose is to explore these properties in a canonical life-cycle model under the assumption that some individuals fail to rationalize the financial incentives embedded in the structure of the pension system.

The introduction or structural reform of earnings-based pension systems is expected to modify incentives for labor supply on both the intensive and extensive margins (e.g., Browning, 1975; Sommacal, 2006; French and Jones, 2012). First, the contribution rate, net of the present value of incremental pension benefits realized from such a transfer, effectively instruments a labor income tax (e.g., Cigno, 2008; Fisher and Keuschnigg, 2010; Buyse et al., 2017). The magnitude of the implicit taxation, in turn, depends on the correlation between individual contributions and entitled benefits. The weaker the link, the higher the implicit tax rate (Cigno, 2008).

[1] For details, see e.g., Palmer (2001)

[2] From a long-run perspective, voters who manifest this type of financial ignorance may also contribute to maintaining or introducing new suboptimal retirement policies.

Second, a pension system is bound to modify the rate at which individuals accumulate a sufficient capital buffer to finance their retirement. For example, a system that redistributes income intragenerationally may allow low- (high-) income households to secure a sufficient retirement income with less (more) labor supplied relative to a system where individual benefits are realized from individual contributions.

This paper argues that these results, to a large extent, follow the assumption that individuals fully comprehend how their labor supply behavior interacts with the benefit formula. Ultimately, if a large number of agents fail to account for how pension systems modify incentives for labor supply over the life cycle, it could have substantial effects on pension system performance in terms of efficiency and equity.

A number of studies support the notion that pension illiteracy is a widespread phenomenon. Elinder et al. (2020) conclude from survey data that a large fraction of Swedes lack basic knowledge of the pension system, with ignorance being especially prominent among young people, women, the less educated, and low-skilled earners. Around 80% of respondents reported the perceived complexity of the pension system as a primary reason for failing to acquire adequate knowledge about its structure. Also using survey data, Chan and Stevens (2008) find that there is a great deal of heterogeneity in how individuals respond to incentives in the U.S. Pension system when choosing their retirement age. Bucher-Koenen et al. (2019) find that most Europeans have inaccurate understanding of their pension systems and that it is primarily well-informed individuals who change their labor supply following pension reforms. Barrett et al. (2015) note that two-thirds of a representative sample of older Irish people failed to understand the fundamental rules of their pension system, which suggests that pension illiteracy is present among individuals of all ages. It is therefore reasonable to assume that a sizable fraction of the population, at least to some degree, fails to rationalize the trade-offs related to the structure of the public pension system.

To study the implications of pension illiteracy on both the intensive and extensive margins of labor, I consider a continuous-time OLG model with a stylized public pension system containing both a redistributive (Beveridgean) and earnings-based (Bismarckian) pillar.[3] Individuals differ in terms of productivity and optimization sophistication. Regarding the latter, I introduce two types of myopic individuals in addition to the canonical rational individual. Type 1 myopes acknowledge that the timing of their retirement affects the annuitization of benefits but fail to recognize how contributions determine benefits. Type 2 myopes treat retirement annuities as strictly exogenous. One can think of these individuals as continually receiving statements of pension wealth from the pension agency. Type 1 myopes receive information about their accumulated pension wealth, and make decisions about how this wealth can be effectively annuitized. Type 2 myopes receive statements on pension wealth in terms of annuities, and thus fully treat benefits as exogenous. As such, the type of information which the pension agency relays to the public matters.

I make no normative assessment of the origin of myopia among individuals and instead assume it is an innate, non-redeemable feature of their behavior in a reduced-form fashion.[4] I further assume that myopes make otherwise fully rational life-cycle plans to ensure that any differences between the behavior of lifecyclers and myopes of identical productivity arise from pension illiteracy only.

The model is analyzed in two ways. In the first step, I analyze formally the optimal control conditions for savings and intensive margin labor supply behavior, as well as optimal retirement age conditions. Key insights related to how implicit taxation induced by the pension system varies with the assumptions of optimization sophistication are obtained.

If the pension system is characterized by a correlation between contributions and realized

[3]While it is common to study the effects of pension system structure on intensive margin labor supply and savings behavior in a Diamond-type model (e.g., Sommacal, 2006; Wen et al., 2015; Frassi et al., 2019), it is difficult to augment such a discrete-time framework to include the retirement margin decision without compromising on analytical tractability.

[4]Many studies on retirement behavior, public pensions, and behavioral deficits model myopia as a phenomenon endogenous to hyperbolic preferences (see e.g., Cremer and Pestieau (2011)), short planning horizons (see e.g., Caliendo and Aadland (2007)), or optimization costs (see e.g., Krusell and Smith Jr (1996)). Without denying that these mechanisms are important for explaining myopia in the context of retirement planning, I do not model myopia as endogenous to a specific behavioral mechanism. Another approach, more common to quantitative studies, is the modeling assumption that some agents do not optimize (see e.g., Caliendo and Findley (2019)). To focus the analysis on the implications of pension illiteracy only, I do not restrict the behavior of myopic agents to that of complete non-optimization.

benefits, lifecyclers will acknowledge the foregone compound interest of contributing to the return-dominated pension system instead of saving privately. The implicit tax rate then becomes relatively higher for younger workers than for older workers. For myopic individuals, the implicit tax rate is constant over the life cycle and independent of the relative weight given to the Bismarckian and Beveridgean pillars.

By assumption, myopic individuals fail to fully account for the incremental pension benefits realized from working for a longer period of time. Therefore, any variations in replacement income will be treated as exogenous. For type 2 myopes, the active-retired trade-off is further simplified by the assumption that they do not strategically annuitize their accumulated pension wealth. As a result, both myope types retire earlier than lifecyclers, with type 2 myopes retiring the earliest. Ultimately, pension illiteracy is found to reduce aggregate labor supply.

To obtain results for how pension illiteracy affects inequality, I turn to numerical simulations. These simulations yield two main results: (1) Pension illiteracy can increase both earnings inequality and lifetime inequality. (2) Because of pension illiteracy, a Beveridgean, flat benefit system can yield higher lifetime inequality relative to a Bismarckian, earnings-based system.

These results suggest that the presence of pension illiteracy makes the equity-efficiency trade-off difficult to characterize. The results reveal that policies that aim to either increase labor market participation by strengthening the link between earnings and benefits, or reduce economic inequality by implementing a redistributive pension system, can have counterintuitive outcomes.

The remainder of the paper is organized as follows. The model is introduced and solved in section 2. Section 3 contains the results. Section 4 concludes.

2 The Model

Consider a continuous-time OLG-type economy in steady state as described in Jacobs (2009) in which individuals of all ages are represented at each instant in time and identically replicate themselves. The economy consists of a continuum of two types of individuals indexed by skill ($i = 1, 2$; $1 =$ low, $2 =$ high). Unless explicitly needed, this indexation will be suppressed to avoid notational clutter. Differences in productivity realized from skill distribution provides a policy motive for redistribution. Individuals identically replicate themselves, and population is held constant and normalized to unity. Λ is the fraction of low-skilled individuals, and residual $1 - \Lambda$ is the fraction of high-skilled individuals. To simplify analysis, I assume fixed and exogenously set factor prices of interest rates r, and wage rates w_i. The economy is small and open, which together with the steady state assumption implies exogenous and fixed factor prices

2.1 Individual utility maximization

By keeping population size constant and assuming identical cohorts, I only need to model the behavior of one generation. The economic life of each individual begins at model age $t = 0$ and ends with certainty at $T > 0$. Individuals derive utility from the consumption of non-durable goods, $c(t) > 0$, time spent on leisure $l(t) \in [0, 1]$, and time spent in retirement (retirement leisure) $T - R_i$, where R_i is the retirement age. During their working life, individuals earn a labor wage w_i for each unit of labor supplied, net of a public pension contribution rate $\tau \geq 0$. During retirement, individuals receive a pension annuity b_i.[5] Any savings flow into the individual's asset account $k(t)$ and earn interest at a risk-free rate $r > 0$. I assume that each individual starts and ends his or her economic life with zero wealth, $k(0) = k(T) = 0$.

Suppose all individuals maximize life-cycle utility as represented by the following additive

[5]Individuals are only eligible for receiving pension income upon full-time retirement. This paper thereby abstracts from the possibility of part-time retiring and the partial receipt of benefits while remaining employed.

separable functional form, subject to the intertemporal budget constraint.:

$$V \equiv \max_{\{c(t),l(t),R\}} \int_0^R [\ln(c(t)) + \beta \ln(l(t))] e^{-\theta t} dt + \int_R^T \ln(c(t)) e^{-\theta t} dt + \eta \frac{(T-R)^{1-\frac{1}{\phi}}}{1-\frac{1}{\phi}}. \tag{1}$$

The asset accumulation equation can be written as:

$$\dot{k} = \begin{cases} (1-l(t))w(1-\tau) + rk(t) - c(t) & \text{for } t \in [0, R) \\ b + rk(t) - c(t) & \text{for } t \in [R, T], \end{cases} \tag{2}$$

In the above optimization problem, $\theta > 0$ is the subjective rate of time preferences, β is the weight attached to leisure during the working life, η is the weight on retirement leisure preferences, and $\phi > 0$ is an elasticity related to the time spent in retirement that makes utility non-linear in retirement leisure. The first integral corresponds to the utility of consumption and leisure during the working life, while the second integral plus the utility derived from retirement leisure corresponds to the utility of the retirement phase.[6] Relative to the baseline model specification in Jacobs (2009), the model used in this study abstracts from the human capital formation while adding heterogeneous individuals in terms of productivity and optimization behavior as well as an endogenous retirement system with both earnings-based and redistributive components. The specification of log utilities in consumption and leisure is a common assumption in the quantitative macro literature and is convenient for the purpose of obtaining analytical expressions.[7] Since the retirement good is defined as years in retirement, I consider the more general CRRA specification since the possible solution of $R = T$ would not be defined under the log specification.

[6]Following Jacobs (2009), the retirement good is not discounted. It is also difficult to justify which age that should be used as a reference in the discounting of such a good since it is a function of the years in retirement but not specified as a flow variable.

[7]The assumption of log utilities follows the specification of the CRRA utility such that the parameters of the intertemporal elasticity of substitution in consumption and leisure are equal to unity. This is reasonable given the empirical evidence reviewed in Thimme (2017) for consumption and Blundell and MaCurdy (1999) for leisure.

2.2 Pension system

This study models a stylized, actuarially unfair PAYG pension system, which aims to resemble the main features of a modern multi-pillar pension system that combines both Bismarckian and Beveridgean transfer mechanisms in the benefit formula.[8] While the Bismarckian and Beveridgean pillars are compulsory, one can consider any private retirement savings to make up the third, voluntary, pillar.

The pension benefit formula for determining individual benefits is specified as follows:

$$b = \frac{\tau}{T-R}\left[\kappa w \int_0^R (1-l(t))dt + (1-\kappa)Y\right], \tag{3}$$

where Y is the aggregate income defined as:

$$Y = \Lambda y_1 + (1-\Lambda)y_2 = \Lambda w_1 \int_0^{R_1} (1-l_1(t))dt + (1-\Lambda)w_2 \int_0^{R_2} (1-l_2(t))dt. \tag{4}$$

$\kappa \in [0,1]$ determines the dependence of realized benefits and individual contributions. If $\kappa = 0$, the system is purely Beveridgean and consists solely of a common benefit level, and if $\kappa = 1$, individual benefits are perfectly correlated with individual contributions and thus replicate a pure Bismarckian system.

The system in equation (3) is actuarially unfair by construction since contributions to the pension system are not capitalized and thus underperform private savings in terms of compound interest. As such, the implicit tax rate induced by the system will always be positive for any $r > 0$.

It is also important to note that this structural representation abstracts from the nonlinear features of the contribution–benefit formula common to real-world systems.[9] This is a common modeling assumption (see e.g., Sommacal, 2006; Cigno, 2008; Caliendo and Findley, 2019) as nonlinear futures typically vary substantially between different institutional settings.

[8]See e.g., World Bank (1994) for a more detailed discussion of such a system design.

[9]For example means-testing, pension penalties and progressivity.

The result in this paper therefore applies to the fundamental transfer mechanisms which guides the overall design of pension systems in most OECD countries. Furthermore, the inclusion of nonlinearities typically makes the structure of the solutions for optimal controls highly untractable (e.g., Wang and Li, 2017).

2.3 Behavioral types

Irrespective of income, individuals can be either lifecyclers or myopes. Lifecyclers behave as canonical rational agents and fully realize that they can affect their pension benefits through their intensive and extensive margin labor supply decisions via the Bismarckian pillar and the annuity divisor, following Equation (3).

The behavior of myopes departs from the behavior of rational agents in one of two ways. While individual pension benefits are always computed based on the formula in Equation (3), type 1 myopes do not account for the value of κ and thus treat the numerator in Equation (3) times the expression in square brackets as exogenous. Thus, they perceive b as follows:

$$b = \frac{B}{T - R}, \tag{3'}$$

where B_i can be thought of as a statement provided by the pension agency on accumulated pension wealth. As a result, the type 1 myope only accounts for the annuity divisor $\left(\frac{1}{T-R_i}\right)$ in the active-retired trade-off.

Type 2 myopes treat their entitled benefit annuities as strictly exogenous and base their decisions on statements of pension income expressed in terms of annuities $\bar{b}$:

$$b = \bar{b}. \tag{3''}$$

By maximizing Equation (1), subject to the asset accumulation equation in Equation (2) and conditional on Equations (3), (3'), or (3") respectively, the optimal consumption and intensive margin leisure trajectories of lifecyclers and myopes can be expressed as follows:

$$c^*(t) = \frac{e^{(r-\theta)t}}{\mu_0} \tag{5}$$

$$l^*(t) = \begin{cases} \frac{\beta e^{-\theta t}}{\mu_0 w[(1-\tau)e^{-rt} + \frac{\kappa\tau}{r(T-R)}(e^{-rR} - e^{-rT})]} & \text{for lifecyclers,} \\ \frac{\beta e^{(r-\theta)t}}{\mu_0 w(1-\tau)} & \text{for type 1 and 2 myopes,} \end{cases} \tag{6}$$

where μ_0 is an unknown constant. Equation (5) illustrates that consumption evolves according to the Euler equation, which holds true for all individuals since myopia is assumed to directly affect only the labor-leisure trade-off. As seen in Equation (6), the intensive margin decision of lifecyclers includes a forward looking effect of labor supply on retirement benefits, while myopes reduce the labor-leisure trade-off to that of only contemporaneous effects. That is, the lifecycler acknowledges that a fraction of contributions is earmarked as future pension income. This effect is, in part, offset by a compound interest effect, as contributions to the pension system are illiquid and therefore return-dominated by private savings. If the risk-free interest rate is zero, and the system is purely Bismarckian, any contributions to the pension system would constitute perfect substitutes to private savings. Under such a scenario, the implicit tax rate would be zero.

The foregone capital income, measured by the compound interest of investing pension contributions in risk-free asset accounts, is determined by the length of the investment horizon. The opportunity cost induced by the Bismarckian pillar will therefore constitute an asymmetric implicit income tax, as the treatment is larger for younger individuals than for older individuals. Note that if $\kappa = 0$ (i.e., if the pension system is purely Beveridgean), the optimality condition for intensive margin leisure is identical for all individuals.

The optimal retirement decision R^* for lifecyclers satisfies the first-order condition that the marginal cost of entering retirement is equal to the marginal cost of delaying retirement:[10]

$$\begin{aligned}&\beta ln(l^*(R^*))e^{-\theta R^*} - \eta[T-R^*]^{-\frac{1}{\phi}} + \mu_0\Bigg\{(1-l(R^*))w\Bigg[(1-\tau)e^{-rR^*}\\&+\frac{\kappa\tau}{T-R^*}\int_{R^*}^{T} e^{-rs}ds\Bigg] - b_i\Bigg[e^{-rR^*} - \frac{1}{T-R^*}\int_{R^*}^{T} e^{-rs}ds\Bigg]\Bigg\} = 0.\end{aligned} \tag{7}$$

The corresponding condition for type 1 myopes take the form:

$$\begin{aligned}&\beta ln(l^*(R^*))e^{-\theta R^*} - \eta[T-R^*]^{-\frac{1}{\phi}} + \mu_0\Bigg\{(1-l(R^*))w(1-\tau)e^{-rR^*}\\&-b_i\Bigg[e^{-rR^*} - \frac{1}{T-R^*}\int_{R^*}^{T} e^{-rs}ds\Bigg]\Bigg\} = 0,\end{aligned} \tag{8}$$

while it takes the following form for type 2 myopes:

$$\beta ln(l^*(R^*))e^{-\theta R^*} - \eta[T-R^*]^{-\frac{1}{\phi}} + \mu_0\Bigg\{(1-l(R^*))w(1-\tau)e^{-rR^*} - b_i e^{-rR^*}\Bigg\} = 0. \tag{9}$$

Some of the components in the active-retired trade-off are identical for all individuals and will therefore not contribute to any differences in retirement behavior between lifecyclers and myopes. In particular, all individuals acknowledge two direct utility effects of delaying retirement. Each individual will gain utility from the fraction of the incremental time of the working life devoted to leisure while experiencing an opportunity cost in terms of foregone utility from retirement leisure.

While the direct effects on utility are homogeneous across individuals, the effects induced by changes to replacement income will differ. These effects are summarized inside the curly brackets in Equations (7)–(9), expressed in present value and converted to utility through the multiplication by the initial marginal utility of wealth μ_0. While all individuals experience an opportunity cost in terms of foregone net labor earnings when they retire, lifecyclers realize that this cost is partly offset by the fraction κ of contributions allocated toward the Bismarckian pillar. Since these contributions are interest-free, the opportunity cost will

[10]See Appendix A for calculations.

never be fully offset if the risk-free interest rate is positive, as any compound interest will constitute a leakage from the lifetime budget. If κ increases, thereby lowering the implicit taxation, the opportunity cost of retiring increases. Ultimately, increasing the relative weight given to the Bismarckian pillar will promote a delayed retirement. Neither type 1 nor type 2 myopes take into account the value of κ and thus experience a lower opportunity cost of retirement in terms of foregone net earnings.

If the benefit level b_i increases, ceteris paribus, the replacement income increases and lowers the cost of retiring. A reform from a Bismarckian system to a Beveridgean system is thus expected to promote an earlier labor market exit among low-skilled individuals as their replacement income increases. Following the same corollary, such a reform is expected to influence high-skilled individuals to delay their retirement. Instead of treating this effect as completely exogenous, both lifecyclers and type 1 myopes acknowledge that the realized annuity amount received at each instant will become larger if they retire at an older age. This effect is, in turn, partly offset by the compound interest that could have been realized by withdrawing pension funds and investing them in a risk-free asset. If the interest rate increases, it becomes financially more attractive to retire earlier. Since type 2 myopes treat the change in b_i as exogenous, they do not recognize the net benefits of delaying retirement.

The main takeaway from Equations (7)–(9) regarding the differences between lifecyclers and myopes can be summarized as follows: (1) Lifecyclers acknowledge that as κ increases, the opportunity cost of retiring increases. This effect is not acknowledged by either myope type. (2) The lifecycler and type 1 myope take into account that their retirement age affects the accumulation of pension wealth via the length of the working life and the annuity divisor. Since both these effects increase the cost of retirement, lifecyclers are expected to retire at an older age relative to type 2 myopes for any value of κ and at an older age relative to type 1 myopes for any $\kappa > 0$. (3) If the system is perfectly Beveridgean, the retirement condition for lifecyclers and type 1 myopes is identical since any pension wealth will be treated as exogenous.

Equations (5)–(9) thereby characterize the optimality conditions for the life-cycle behavior of rational and myopic individuals, respectively.

3 Simulations

3.1 Population structure

I assume that half of the population consists of low-skilled, and the residual half are high-skilled (i.e., $\Lambda = 0.5$ (e.g., Golosov et al., 2013)). A valid concern regards how pension illiteracy is correlated to the income level of the individual.[11] As Elinder et al. (2020) suggest that pension illiteracy is overrepresented among low-skilled individuals, I consider five scenarios for the simulations: (1) a benchmark economy where all individuals are lifecyclers, (2) all individuals are type 1 myopes, (3) all individuals are type 2 myopes, (4) high-skilled individuals are lifecyclers, and low-skilled individuals are type 1 myopes, and (5) high-skilled individuals are lifecyclers, and low-skilled individuals are type 2 myopes.[12]

3.2 Parametrization

I set $T = 55$, which implies that an agent who enters economic life at age 25 lives with certainty to age 80. Since the fundamental analyses concern the implications of pension design under different behavioral assumptions, I calibrate the model to a benchmark scenario of self-financing agents $\tau = 0$. Simulations will subsequently illustrate how the inclusion of public pension modifies behavior relative to a scenario without public pension. The wage rates are set to $w_1 = 1$ and $w_2 = 1.7$, such that the wage (human capital) premium is consistent with that used in e.g., Acemoglu (2002), Sommacal (2006), and Hachon (2010). The risk-free interest rate is set to $r = 3.5\%$. The public pension contribution rate τ is set to

[11] Myopia is commonly introduced into heterogeneous agent models to quantitatively account for the fraction of low-skilled individuals who live hand to mouth as opposed to a phenomenon displayed over the entire income distribution.

[12] The numerical solver is described in Appendix B. Programming was conducted in SageMath, a computer algebra system with a Python-like syntax. All codes are available upon request from the author.

vary between 0 and 30 % to encompass the public pension contribution rates of most OECD countries.

To reconcile the model output with several stylized facts representative of a small OECD country, I specify the following targets for the benchmark simulation: (i) Decreasing labor market participation over the life cycle. This target limits the subjective discount rate to values below the risk-free interest rate ($\theta < r$). I set $\theta = 2\%$. (ii) Average retirement age is 63–65. (iii) Average weekly working hours are close to 35.[13] To simultaneously achieve targets (ii) and (iii), conditional on θ, the following parametrization was used: $\beta = 1.5$, $\eta = 0.6$, and $\phi = 0.9$.

3.3 Life-cycle labor supply

3.3.1 Changing the structure of public pensions

I begin by considering a parametric reform of the public pension system for a given value of the contribution rate. Fixing $\tau = 0.15$, I proceed to vary $\kappa \in [0, 1]$. As previously discussed, $\kappa = 0$ implies a purely Beveridgean pension system in which all contributions are allocated toward a flat-rate benefit pillar. The opposite corner of $\kappa = 1$ corresponds to a purely Bismarckian system, implying zero intragenerational redistribution via the benefit formula. Figure (1) illustrates how individual labor supply changes with κ, while Figure (2) compares changes in aggregate labor supply for the different assumptions on population structure. Note that the figure depicts the efficient labor supply of individuals by scaling labor supply by the human capital (wage) premium.

To facilitate the interpretation of the numerical result, recall some of the analytical insights obtained in section 2. From the cases described in Equation (6), it is clear that a change in κ will only modify the intensive margin labor-leisure trade-off for lifecyclers. Myopic individuals will instead treat contributions, irrespective of the value of κ, as a labor

[13] I assume that the total time available in the model for allocating between work and leisure corresponds to 5 days per week and 17 hours per day. This is consistent with the time endowment of 84 hours per week used in Goulder et al. (2019).

income tax, which will lower the perceived payoff at each instant throughout the working life $t \in [0, R)$. Since the specification of log utilities implies that income and substitution effects perfectly offset each other, such proportional tax treatment will not affect the optimal time allocation between intensive margin labor supply and leisure. As such, changing the relative weights of the Bismarckian and Beveridgean pillars will not directly affect myopes' intensive margin labor supply decisions. Lifecyclers will, however, realize that increasing κ lowers the implicit taxation of the contribution rate, as any contributions to the Bismarckian pillar will be realized as future benefits. The implicit tax rate will always be positive as long as $r > 0$, since contributions to the Bismarckian pillar are illiquid and thus return-dominated by private savings. In particular, the implicit taxation is relatively higher for young workers as the foregone interest of any pension contributions decreases with the time left until retirement. As a result, this asymmetric tax treatment makes labor supply relatively more (less) attractive when workers are old (young), which influences the individual to substitute labor supply from a young age to an older age. The lower the risk-free interest rate, the lower the implicit taxation following the foregone compound interest.

Regarding the extensive margin decision, Equations (8) and (9) inform that modifying κ will only affect the retirement decision of myopes via changes to the statements on replacement income. For example, Equation (9) implies that the optimal retirement age of type 2 myopes coincides with the point in time where the utility of retirement leisure and retirement benefits compensates for the foregone utility of working life leisure and earnings. Reforming the pension system from Beveridgean to Bismarckian introduces opposite effects for high and low-skilled: A low earner will face a lower replacement income and thus experience an increase in the cost of retirement. As a result, she will delay retirement. The opposite will hold true for high-skilled who will retire earlier following an increase in replacement income, as a larger share of contributions are realized as future benefits. As evidenced in Figures (1)–(2), these labor supply effects perfectly offset each other, suggesting that a parametric reform toward a Bismarckian system has no effect on aggregate labor supply if agents are myopic. If individuals behave as type 2 myopes, the difference is an overall lower degree of labor market participation following an inability to effectively annuitize pension wealth.

From Equations (6)–(8), it can be shown that when $\kappa = 0$, the optimization problems of the lifecycler and type 1 myope are identical. This is intuitive since accumulated pension wealth under a pure Beveridgean system is viewed as exogenous not only by the type 1 myope but also by the lifecycler, given zero correlation between contributions and benefits. As a result, both agents only take into account the annuitization of pension wealth when deciding on their retirement age. This is illustrated in Figure (1). As κ increases, however, both high- and low-earner lifecyclers will acknowledge that a larger share of their contributions is allocated toward their individual pension accounts. Since the Bismarckian pillar constitutes a less beneficial savings mechanism relative to risk-free funds, the net payoff to labor supply decreases and thus lowers the opportunity cost of retiring. Since the high (low) earner experiences a lower (higher) replacement income following such a reform, the labor supply experiences a disproportionately greater increase among low-skilled relative to high-skilled. As illustrated in Figure (2), lifecyclers and type 1 myopes behave identically if $\kappa = 0$. The aggregate labor supply effects in the economy will then be identical in a scenario where the population is composed of high-earner lifecyclers and low-earner type 1 myopes, similar to if the population was homogeneous, with only lifecyclers or type 1 myopes. When increasing the relative weight of the Bismarckian pillar, the low-skilled experience increased incentives for delaying retirement as her realized pension wealth decreases following a lower degree of redistribution. However, since the low-earner myope does not account for the lower relative implicit taxation she will receive when older, the increase in labor supply is not as large as if she had been a lifecycler. As such, aggregate labor supply increases, but not as much as if both individuals had behaved as lifecyclers. The result is qualitatively the same when assuming that the population consists of high-earner lifecyclers and low-earner type 2 myopes, except that the labor supply of the low-skilled household is generally lower.

3.3.2 Changing the contribution rate

In this section, I study the labor supply effects by varying the contribution rate to a pure Bismarckian ($\kappa = 1$) and pure Beveridgean ($\kappa = 0$) system for the different assumptions of population structure. Figure (3) illustrates individual responses to a change in the contribution rate, and Figure (4) shows the aggregate effects.

The main intuition carries over from the last subsection. When considering a scenario with only lifecyclers in a Beveridgean system, increasing the contribution rate will result in a greater reduction in aggregate labor supply. By introducing a pure Beveridgean system, two effects promote an earlier retirement. For both high and low-skilled, the pension system lowers the cost of retirement following the introduction of pension benefits. At the same time, the implicit taxation of the contribution rate lowers the opportunity cost of retiring in terms of foregone earnings. As income is redistributed from high to low-skilled, the incentives are greater for the low-earner individual to retire earlier. Since the accumulation of pension wealth is treated as fully exogenous under a Beveridgean system, the labor supply distortions will be of equal size for a scenario with high- and low-earner lifecyclers and a scenario with only type 1 myopes.

The introduction of a Bismarckian system implies that income is redistributed only over the life cycle as opposed to from high to low-skilled. In a scenario with only lifecyclers, increasing the size of a Bismarckian system will therefore not generate disincentives for labor supply to the extent of a Beveridgean system. Figure (4) illustrates that if both households are type 1 myopes, increasing the size of the Bismarckian system results in aggregate labor supply effects that are identical to increasing the size of a Beveridgean system with either lifecyclers or type 1 myopes. This follows from the fact that individuals in both scenarios view accumulated pension wealth as exogenous. Thus, the only effect of increasing the contribution rate in both scenarios is in terms of changes to the replacement income. A high- (low-) earner lifecycler under a Bismarckian system will experience less (more) of a loss in terms of replacement income when increasing the contribution rate relative to what a high- (low-) earner lifecycler or type 1 myope would experience under a Beveridgean system. Consistent with the result illustrated in Figure (2), these differences in changes to labor supply among high and low-skilled offset each other. The results are qualitatively the same when comparing the labor supply response of type 2 myopes under a Bismarckian and Beveridgean system, although the magnitudes of labor supply distortions are larger following the inability to comprehend the annuity divisor.

From Figure (4), it is clear that when the high-skilled individual is a lifecycler, and the low-skilled individual is a myope of any kind, increasing the contribution rate in a Beveridgean system leads to a larger decrease in labor supply when compared to increasing the rate in a Bismarckian system. This is reasonable since the high earner under a Bismarckian system rationalizes all her contributions as future income, while she would treat contributions as a tax under a Beveridgean system. Since the low-skilled individual suffering from myopia always treats the contribution rate as a labor income tax, the aggregate decrease in labor supply will be larger if the system is Beveridgean.

3.4 Intragenerational inequality

Since there are only two productivity types in each simulation, the income inequality measures are constructed as the ratio of income types expressed in present value terms between high- and low-skilled individuals, as follows:

$$\text{Earnings inequality (EI)} \;=\; \frac{(1-\Lambda)w_2\int_0^{R_2}(1-l_2(t))e^{-\gamma t}dt}{\Lambda w_1\int_0^{R_1}(1-l_1(t))e^{-\gamma t}dt}, \tag{10}$$

$$\text{Pension inequality (LI)} \;=\; \frac{(1-\Lambda)\int_{R_2}^{T}b_2e^{-\gamma t}dt}{\Lambda\int_{R_1}^{T}b_1e^{-\gamma t}dt}, \tag{11}$$

$$\text{Lifetime inequality (LI)} \;=\; \frac{(1-\Lambda)(w_2\int_0^{R_2}(1-l_2(t))e^{-\gamma t}dt+\int_{R_2}^{T}b_2e^{-\gamma t}dt)}{\Lambda(w_1\int_0^{R_1}(1-l_1(t))e^{-\gamma t}dt+\int_{R_1}^{T}b_1e^{-\gamma t}dt)}. \tag{12}$$

In Equations (10)–(12), γ denotes an arbitrary discount rate used by the policymaker to compute the present value of future income. Following Sommacal (2006), I set $\gamma = 0$, which implies that both labor and pension income are given the same weight in the measure of lifetime inequality.[14] Figure (5) illustrates how earnings inequality varies with the contribution rate, Figure (6) shows pension inequality, and Figure (7) presents lifetime inequality. Since the wage premium is fixed, the dispersion in lifetime labor supply constitutes the only variable influencing inequality. The results in section 3.3 are therefore fundamental for understanding the realized effects in terms of earnings, pensions, and lifetime inequality. As previously discussed, and in agreement with the findings in Sommacal (2006), the Beveridgean system increases the difference in lifetime labor supply between high- and low-skilled individuals. In comparison to a Bismarckian system, the low earner receives higher benefits and thus experiences an additional financial incentive for retiring earlier. As such, the Beveridgean system increases inequality in the lifetime labor supply, which translates to an increase in earnings inequality, as illustrated in Figure (5).

Increasing the size of the Bismarckian system will increase earnings inequality only in the scenarios with high-skilled lifecyclers and low-skilled myopes. When considering mixed regimes with high-skilled lifecyclers and low-skilled myopes of any type, the dispersion in effective labor supply will not be proportional to the wage premium, as the low earner reduces her labor supply to a disproportionately greater extent following an increase in the contribution rate. The effect of the different pension structures on pension inequality is illustrated in Figure (6). Since benefits under the Bismarckian system are proportional to the individual's earnings history, the realized pension inequality will be equal to the wage premium of 1.7 as long as both productivity types are lifecyclers.

Simulations illustrate the two opposite effects on inequality induced by a Beveridgean pension system. On the one hand, the system redistributes pension wealth such that pension inequality decreases. On the other hand, the system promotes increased earnings inequal-

[14]See the first section of Hancock and Richardson (1985) and references therein for an elaborate discussion about the choice of discount rate when computing a value for income inequality.

ity. Concluding which of these effects dominates is not straightforward. In particular, by increasing the contribution rate, a greater share of lifetime income will be realized as pension income, and a lesser share will be realized as earnings. The higher the contribution rate, the lower the net earnings will be, and as a result, less weight will be attributed to earnings inequality in the lifetime inequality measure.

Figure (7) illustrates the total effect on lifetime inequality by varying the contribution rate under different structural assumptions of the pension system. For scenarios with homogeneous populations, lifetime inequality will be equal to the human capital premium for any size of the contribution rate as long as the system is perfectly Bismarckian. If the system is purely Beveridgean and the population consists of lifecyclers or type 1 myopes, the redistributive effect on pension income will dominate the increased earnings inequality.

Results obtained thus far conform well with the equity-efficiency trade-off: The Beveridgean system reduces overall economic inequality but results in less aggregate labor supply.

When considering a scenario with only type 2 myopes, however, a stand-out result is obtained: The increased earnings inequality is found to dominate the reduced pension inequality. Nevertheless, as the contribution rate increases beyond a value of 0.1–0.15, the realized lifetime inequality begins to decrease. This finding can be attributed to the redistribution of income over the life cycle. By increasing the contribution rate, a larger fraction of income is realized as pension income while net earnings decrease. This implies that earnings are given a smaller relative weight in the measurement of lifetime inequality and, subsequently, in the inequality measure. However, for the range of contribution rates considered in this study, the Beveridgean system always results in higher lifetime inequality in comparison to a Bismarckian system where all individuals are type 2 myopes.

Since the discount rate used to compute the present value of lifetime income is set to zero, the policymaker gives equal weight to earnings and pension income in the inequality measure. However, if using other commonly implemented discount rates, such as the interest rate or the rate of time preferences, the increased inequality imposed by the Beveridgean system would be amplified, as a larger relative weight would be given to the dispersion in earnings in the inequality measure. This ultimately suggests that the presence of pension illiteracy could, if widespread, introduce a substantial ambiguity regarding the equity-efficiency trade-off between a Bismarckian and Beveridgean system.

Observing the results for heterogeneous populations adds some nuance to this finding. Since both pension and earnings inequality increases under a Bismarckian system with high-earner lifecyclers and low-earner myopes, improvements can be made in terms of equity when reforming the pension system from Bismarckian to Beveridgean. Since the Beveridgean scenario with high-skilled lifecyclers and type 1 myopes will yield the same outcome in terms of labor supply as if both individuals were lifecyclers, the policymaker could achieve lower lifetime inequality by increasing the size of the pension system. For any other scenario considered for the simulations, however, increasing the size of the pension system will lead to an increase in lifetime inequality. In the Bismarckian scenarios with high-skilled lifecyclers and low-skilled myopes, the increased difference in labor supply induced by the pension system will substantially increase inequality.

4 Concluding Remarks

This paper models the effects of pension illiteracy on aggregate labor supply and its implications for the redistributive performance of public pension design. I introduce a combined Bismarckian-Beveridgean pension system into an OLG model in the spirit of Jacobs (2009). Individuals are assumed to differ in productivity and pension literacy: lifecyclers adhere to the rational agents paradigm and therefore fully acknowledge the economic incentives embedded in public pensions when planning life-cycle labor supply. Type 1 myopes do not rationalize the contribution–benefit formula but acknowledge that their retirement decision affects annuities through the annuity divisor. Type 2 myopes treat benefit annuities as strictly exogenous. The model is solved as a two-stage delayed response problem, which makes it possible to obtain optimality conditions via Pontryagin's maximum principle.

Results are obtained both analytically and through numerical simulations: (1) Myopia can reduce life-cycle labor supply, with type 2 myopes displaying a substantially lower degree of labor market participation. (2) Pension illiteracy can increase earnings and lifetime income inequality. (3) When type 2 myopia dominates across income groups, a Beveridgean public pension system can increase intragenerational inequality. These findings imply two important policy conclusions. First, if pension illiteracy is a widespread phenomenon, current movements toward more earnings-based pension systems may not increase labor market participation since myopes continue to treat contributions to the pension scheme as a labor income tax. Second, it is not obvious that a reform which increases the relative weight of the Beveridgean pillar will reduce lifetime inequality if pension illiteracy is prominent. These findings suggest that pension illiteracy introduces a problematic ambiguity for policymakers aiming to resolve the equity-efficiency trade-off, and informing the public about the features of the public pension system may promote both efficiency and equity gains.

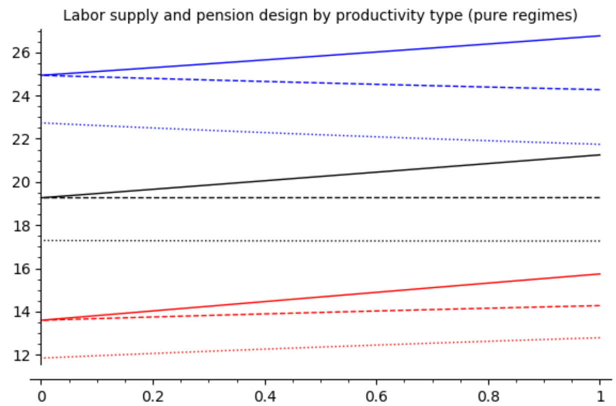
This paper models pension illiteracy as an exogenously imposed characteristic. While this approach is analytically convenient, as it allows the researcher to obtain closed-form solutions for most optimality conditions, it is a limitation from a theoretical viewpoint. A potential approach to endogenize the behavioral failure would be to include various costs of acquiring adequate information about the pension system as an optimization cost. Such a model could then be used to study the quantitative impacts of government information campaigns aimed at reducing the cost of searching for and selecting information for individuals regarding their pensions. In addition, an important question for future modeling is to what degree ignorance is inherited. In this paper, it is treated as entirely inherited, following an assumption of identical replication of individuals. To make a correct assessment of population structure, more empirical evidence is needed.

Lastly, while some key analytical insights are obtained on behalf of the highly tractable solution structure of optimal labor supply behavior, many of the results in this paper are nevertheless obtained through numerical simulations. I therefore end the paper by stressing the importance of more analytical work on the equity-efficiency trade-off when both the intensive and extensive margins of labor supply are endogenous. During times when pension reform is viewed as paramount for hedging against the fiscal stress induced by aging populations, such insights are valuable.

Figures

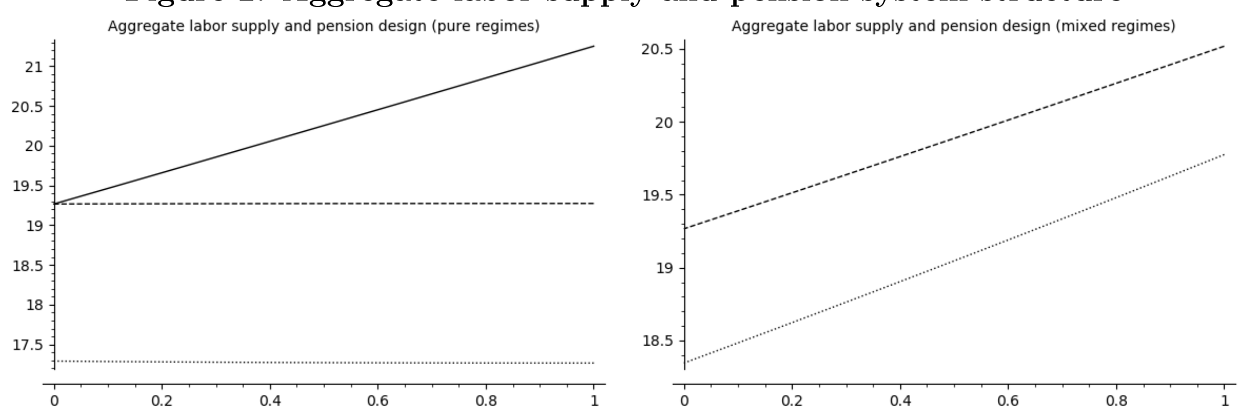
Labor supply

Figure 1: Individual labor supply and pension system structure



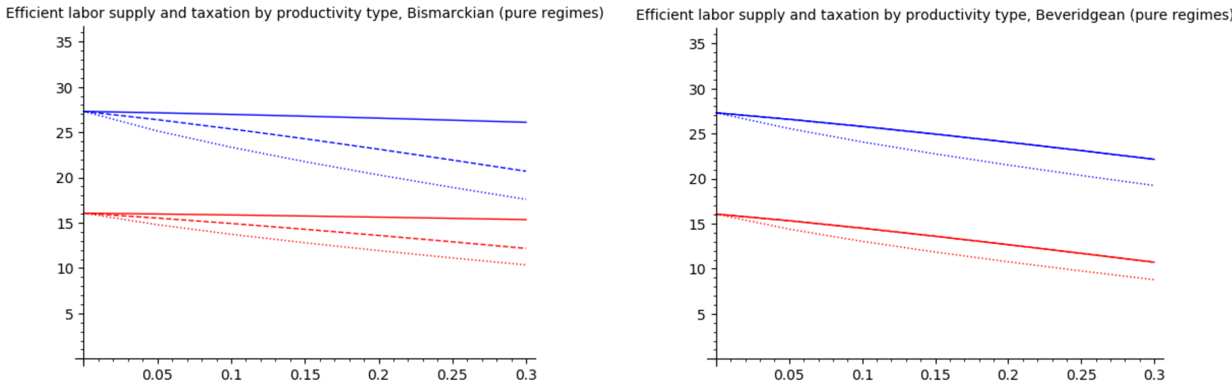
Note: Individual lifetime labor supply (y-axis) for different values of κ *(x-axis), given* $\tau = 0.15$*. Blue lines correspond to high-skilled and red lines to low-skilled. Black represents the weighted average (aggregate) labor supply. Solid lines represent lifecyclers. Dashed lines represent type 1 myopes. Dotted lines represent type 2 myopes.*

Figure 2: Aggregate labor supply and pension system structure



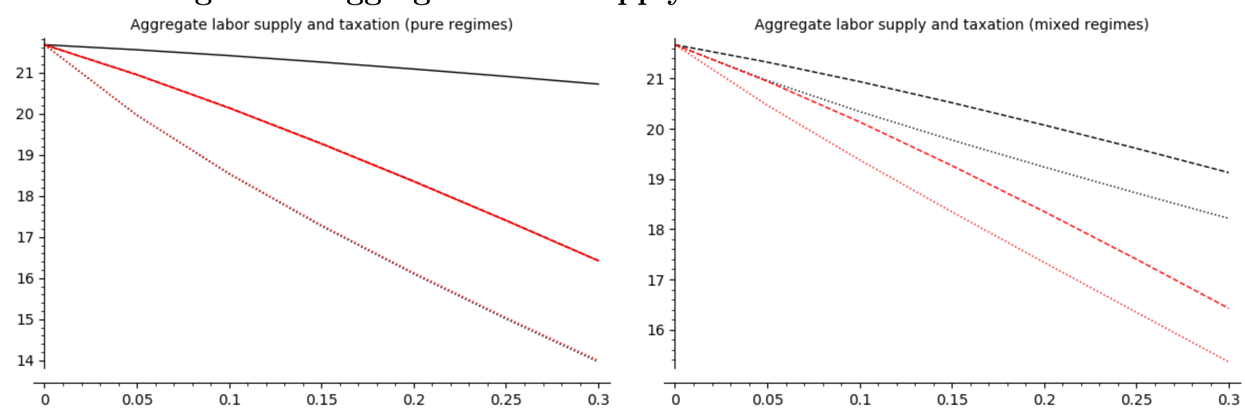
Note: aggregate labor supply (y-axis) for different values of κ *(x-axis), given* $\tau = 0.15$*. Left-hand-side figure illustrates scenarios with homogeneous populations: solid lines represent lifecyclers, dashed lines type 1 myopes, and dotted lines type 2 myopes. The RHS figure illustrates mixed populations: dashed lines represent high-skilled lifecyclers + low-skilled type 1 myopes. Dotted lines represent high-skilled lifecyclers + low-skilled type 2 myopes.*

Figure 3: Individual labor supply and the contribution rate



Note: individual lifetime labor supply (y-axis) for different values of τ (x-axis). The LHS illustrates behavior under a Bismarckian system. The RHS figure illustrates behavior under a Beveridgean system. Blue lines correspond to high-skilled and red lines to low-skilled. Solid lines represent lifecyclers, dashed lines type 1 myopes, and dotted lines type 2 myopes.

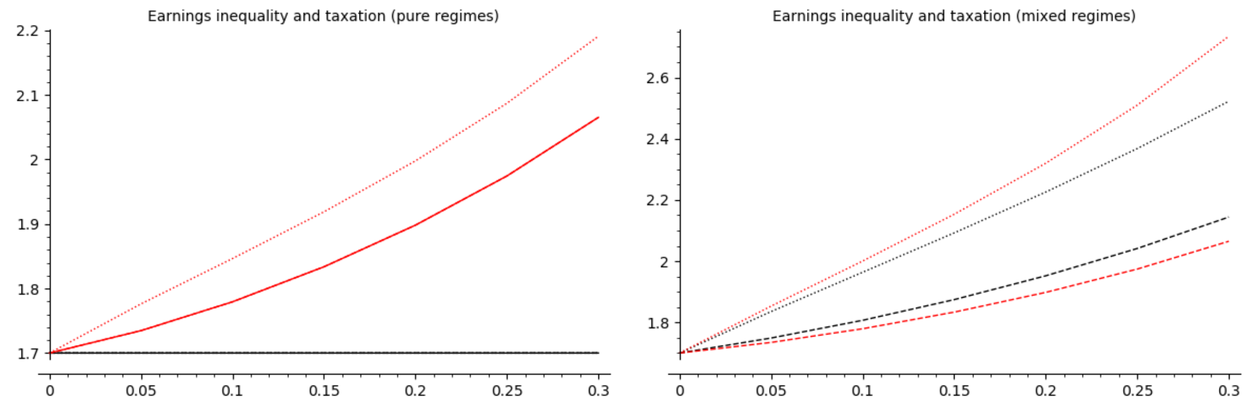
Figure 4: Aggregate labor supply and the contribution rate



Note: aggregate labor supply (y-axis) for different values of τ (x-axis). Black lines corresponds to aggregate labor supply under a Bismarckian system and red lines to under a Beveridgean system. The LHS illustrates scenarios with homogeneous populations: solid lines represent lifecyclers, dashed lines type 1 myopes, and dotted lines type 2 myopes. RHS illustrates the scenarios with mixed populations: dashed lines represent high-skilled lifecyclers and low-skilled type 1 myopes. Dotted lines represent high-skilled lifecyclers and low-skilled type 2 myopes.

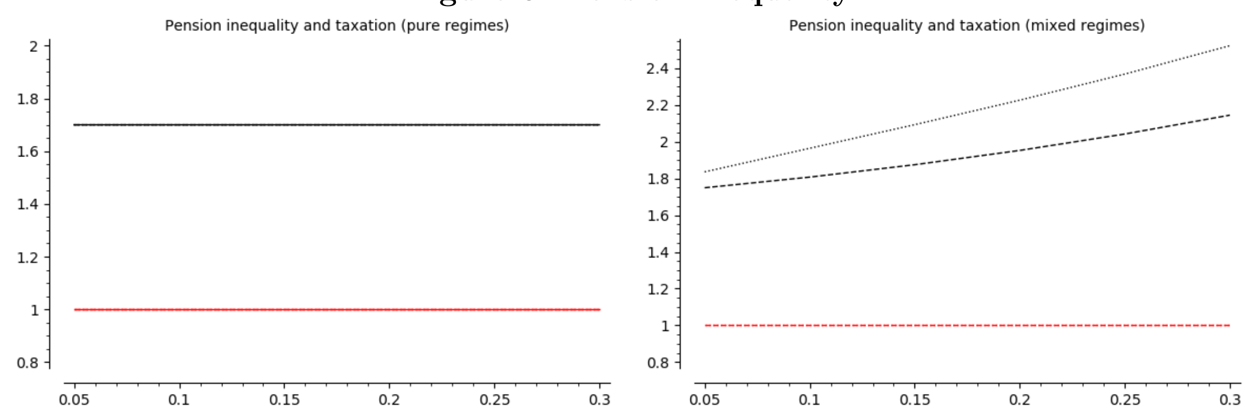
Inequality

Figure 5: Earnings inequality



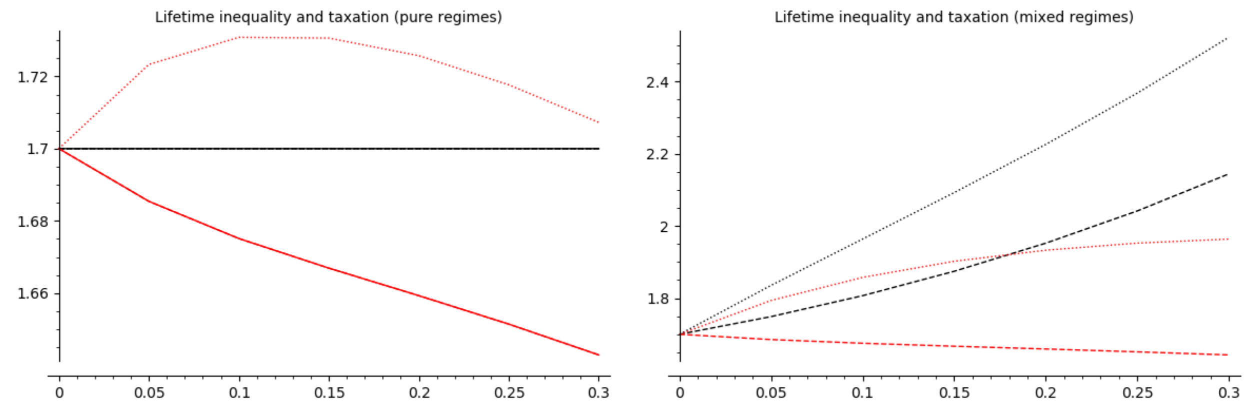
Note: earnings inequality (y-axis) for different values of τ (x-axis). Black lines illustrates Bismarckian scenarios, and red lines Beveridgean scenarios. The LHS illustrates scenarios with homogeneous populations: solid lines represent lifecyclers, dashed lines type 1 myopes, and dotted lines type 2 myopes. The RHS illustrates scenarios with heterogeneous populations: dashed line represent high-skilled lifecyclers + low-skilled type 1 myopes. Dotted lines represent high-skilled lifecyclers + low-skilled type 2 myopes.

Figure 6: Pension inequality



Note: pension inequality (y-axis) for different values of τ (x-axis). Black lines illustrates Bismarckian scenarios, and red lines Beveridgean scenarios. The LHS illustrates scenarios with homogeneous populations: solid lines represent lifecyclers, dashed lines type 1 myopes, and dotted lines type 2 myopes. The RHS illustrates scenarios with heterogeneous populations: dashed line represent high-skilled lifecyclers + low-skilled type 1 myopes. Dotted lines represent high-skilled lifecyclers + low-skilled type 2 myopes.

Figure 7: Lifetime inequality



Note: lifetime inequality (y-axis) for different values of τ (x-axis). Black lines illustrates Bismarckian scenarios, and red lines Beveridgean scenarios. The LHS illustrates scenarios with homogeneous populations: solid lines represent lifecyclers, dashed lines type 1 myopes, and dotted lines type 2 myopes. The RHS illustrates scenarios with heterogeneous populations: dashed line represent high-skilled lifecyclers + low-skilled type 1 myopes. Dotted lines represent high-skilled lifecyclers + low-skilled type 2 myopes.

References

Acemoglu, D. (2002). Technical change, inequality, and the labor market. *Journal of economic literature*, 40(1):7–72.

Barrett, A., Mosca, I., and Whelan, B. (2015). How well-informed are pension scheme members on their future pension benefits? evidence from ireland. *Journal of aging & social policy*, 27(4):295–313.

Blundell, R. and MaCurdy, T. (1999). Labor supply: A review of alternative approaches. *Handbook of labor economics*, 3:1559–1695.

Browning, E. K. (1975). Labor supply distortions of social security. *Southern Economic Journal*, pages 243–252.

Bucher-Koenen, T., Ferrari, I., and Pettinicchi, Y. (2019). The role of pension knowledge in labour supply decisions: Evidence from europe. Technical report, Netspar Discussion Paper 11/2019-043.

Buyse, T., Heylen, F., and Van de Kerckhove, R. (2017). Pension reform in an olg model with heterogeneous abilities. *Journal of Pension Economics & Finance*, 16(2):144–172.

Caliendo, F. and Aadland, D. (2007). Short-term planning and the life-cycle consumption puzzle. *Journal of Economic Dynamics and Control*, 31(4):1392–1415.

Caliendo, F. N. and Findley, T. S. (2019). Myopia, education, and social security. *International Tax and Public Finance*, pages 1–27.

Chan, S. and Stevens, A. H. (2008). What you don't know can't help you: Pension knowledge and retirement decision-making. *The Review of Economics and Statistics*, 90(2):253–266.

Cigno, A. (2008). Is there a social security tax wedge? *Labour Economics*, 15(1):68–77.

Cremer, H. and Pestieau, P. (2011). Myopia, redistribution and pensions. *European Economic Review*, 55(2):165–175.

Elinder, M., Hagen, J., Nordin, M., and Säve-Söderbergh, J. (2020). Svenska folkets kunskaper om pensionen.

Fisher, W. H. and Keuschnigg, C. (2010). Pension reform and labor market incentives. *Journal of Population Economics*, 23(2):769–803.

Frassi, B., Gnecco, G., Pammolli, F., and Wen, X. (2019). Intragenerational redistribution in a funded pension system. *Journal of Pension Economics & Finance*, 18(2):271–303.

French, E. and Jones, J. (2012). Public pensions and labor supply over the life cycle. *International Tax and Public Finance*, 19(2):268–287.

Golosov, M., Troshkin, M., Tsyvinski, A., and Weinzierl, M. (2013). Preference heterogeneity and optimal capital income taxation. *Journal of Public Economics*, 97:160–175.

Goulder, L. H., Hafstead, M. A., Kim, G., and Long, X. (2019). Impacts of a carbon tax across us household income groups: What are the equity-efficiency trade-offs? *Journal of Public Economics*, 175:44–64.

Hachon, C. (2010). Do beveridgian pension systems increase growth? *Journal of Population Economics*, 23(2):825–831.

Hancock, K. and Richardson, S. (1985). Discount rates and the distribution of lifetime earnings. *Journal of Human Resources*, pages 346–360.

Jacobs, B. (2009). Is prescott right? welfare state policies and the incentives to work, learn, and retire. *International Tax and Public Finance*, 16(2):253.

Kamien, M. I. and Schwartz, N. L. (2012). *Dynamic optimization: the calculus of variations and optimal control in economics and management*. Courier Corporation.

Krusell, P. and Smith Jr, A. A. (1996). Rules of thumb in macroeconomic equilibrium a quantitative analysis. *Journal of Economic Dynamics and Control*, 20(4):527–558.

Palmer, E. (2001). The new swedish pension system.

Pensionsmyndigheten (2020). Vanliga missuppfattningar om pensionen: Pensionsmyndigheten: Pensionsmyndigheten.

Sommacal, A. (2006). Pension systems and intragenenerational redistribution when labor supply is endogenous. *Oxford Economic Papers*, 58(3):379–406.

Thimme, J. (2017). Intertemporal substitution in consumption: A literature review. *Journal of Economic Surveys*, 31(1):226–257.

Wang, Z. and Li, Y. (2017). An indirect method for inequality constrained optimal control problems. *IFAC-PapersOnLine*, 50(1):4070–4075.

Wen, X., Pammolli, F., and Gnecco, G. (2015). Labor supply distortion and capital accumulation when public pension systems differ in an aging economy.

World Bank (1994). *Averting the old age crisis: Policies to protect the old and promote growth. Summary*. The World Bank.

Appendix A. Technical appendix

This appendix contains the step-by-step calculations to solve the optimization problem of the lifecyclers. Recall the maximand:

$$\max_{\{c(t),l(t),R\}} \int_0^R [ln(c(t)) + \beta ln(l(t))]e^{-\theta t}dt + \int_R^T ln(c(s))e^{-\theta s}ds + \frac{\eta[T-R]^{1-\frac{1}{\phi}}}{1-\frac{1}{\phi}}. \tag{13}$$

The reason for introducing a separate time variable s is for notational clarity when considering that the leisure choice during any point in time during the working life will affect the realized pension income annuities throughout retirement. The dynamics of savings can be effectively decomposed into two distinct equations as follows:

$$\dot{k} = \begin{cases} (1-l(t))w(1-\tau) + rk(t) - c(t) & \text{for } t \in [0, R), \\ b(R, l(t)) + rk(s) - c(s) & \text{for } s \in [R, T]. \end{cases} \tag{14}$$

Maximizing Equation (13) subject to Equation (14) can be thought of as a *free switching point two-stage delayed response control problem.* The distinct change to the maximand and the asset accumulation function at time R denotes the terminus of the first stage control problem and the initiation of the second stage control problem. Since R is a decision variable, it constitutes a free switching point. The problem is characterized by a delayed response in the leisure choice during the working life as it has an effect on realized pension income at a latter stage of the lifecycle.

I begin by defining the two Hamiltonian functions corresponding to each stage of the optimization problem:

$$\mathcal{H}_1(t) = [ln(c(t)) + \beta ln(l(t))]e^{-\theta t} + \mu_1(t)[(1-l(t))w(1-\tau) + rk(t) - c(t)], \tag{15}$$

$$\mathcal{H}_2(s, R) = ln(c(s))e^{-\theta s} + \mu_2(s)[b(R, l(t)) + rk(s) - c(s)]. \tag{16}$$

The equations characterizing the optimal consumption and leisure behavior are obtained through the maximum principle. For the first time domain, $t \in [0, R)$:

$$\frac{\partial \mathcal{H}_1(t)}{\partial c(t)} = \frac{e^{-\theta t}}{c(t)} - \mu_1(t) = 0. \tag{17}$$

The effect of a small change in intensive margin leisure is partially realized contemporaneously through the instantaneous labor-leisure tradeoff of leisure utility and foregone labor income, but also partially in a forward-looking fashion as it affect future pension benefits. The maximum principle thus has to be modified to account for this delayed response as follows:

$$\frac{\partial \mathcal{H}_1(t)}{\partial l(t)} + \int_R^T \frac{\partial \mathcal{H}_2(s, R, l(t))}{\partial l(t)}ds = \frac{\beta e^{-\theta t}}{l(t)} - \mu_1(t)w(1-\tau) + \int_R^T \mu_2(s)\frac{\partial b(R, l(t))}{\partial l(t)}ds = 0, \tag{18}$$

$$\dot{\mu}_1 = -r\mu_1(t). \tag{19}$$

For the second time domain, $s \in [R, T]$:

$$\frac{\partial \mathcal{H}_2(s)}{\partial c(s)} = \frac{e^{-\theta s}}{c(s)} - \mu_2(s) = 0, \tag{20}$$

$$\dot{\mu}_2 = -r\mu_2(s). \tag{21}$$

Solving the differential equations in Equations (19) and (21) yield the following expressions for the law of motions:

$$\mu_1(t) = \mu_1(0)e^{-rt} = \mu_0 e^{-rt}, \tag{22}$$

$$\mu_2(s) = \mu_2(R)e^{-r(s-R)}. \tag{23}$$

Implementing the transversality condition that $\mu_1(R) = \mu_2(R)$ (see e.g., Kamien and Schwartz (2012)), Equation (20) can be rewritten as follows:

$$\mu_2(s) = \mu_1(R)e^{-r(s-R)} = \mu_0 e^{-rR}e^{-r(s-R)} = \mu_0 e^{-rs}. \tag{24}$$

Realizing that s is simply a continuation of the time continuum beyond time $t = R$, Equation (24) can in turn be rewritten as:

$$\mu_1(t) = \mu_0 e^{-rt}. \tag{25}$$

Substituting Equation (25) into Equations (17) and (20) thus allows me to obtain the following expression for optimal consumption:

$$c^*(t) = \frac{e^{(r-\theta)t}}{\mu_0}, \tag{26}$$

which implies that the dynamics of optimal consumption follows the conventional Euler equation. By substituting Equation (25) into Equation (18), the expression for optimal leisure can be obtained:

$$l^*(t) = \frac{\beta e^{-\theta t}}{\mu_0 w[(1-\tau)e^{-rt} + \frac{\kappa\tau}{T-R}\int_R^T e^{-rs}ds]}. \tag{27}$$

Equation (27) implies that the optimal leisure profile follows from the tradeoff between utility of non-working and foregone labor and pension income. Since pension contributions does not yield any interest following the PAYG setup with zero population growth, there is a leakage from the life-cycle budget constraint following participation in public pension.

The final step in solving the model is to obtain the condition for optimal retirement timing. I obtain this by first substituting the optimal controls in Equations (26) and (27) into the maximand in Equation (13). For the purpose of further mathematical operations, I rewrite the maximand in terms of its Hamiltonian functions:

$$\begin{aligned} V = &\int_0^R [\mathcal{H}_1(t, c^*(t), l^*(t), k^*(t), \mu_1(t)) - \mu_1\dot{k}]dt \\ &+ \int_R^T [\mathcal{H}_2(s, R, c^*(s), k^*(s), \mu_2(s)) - \mu_2\dot{k}]ds + \frac{\eta[T-R]^{1-\frac{1}{\phi}}}{1-\frac{1}{\phi}}. \end{aligned} \tag{28}$$

By integration-by-parts, Equation (28) can be rewritten as:

$$
\begin{aligned}
V = &\int_0^R [\mathcal{H}_1(t, c^*(t), l^*(t), k^*(t), \mu_1(t)) - \dot{\mu}_1 k(t)]dt + \mu_0 k(0) - \mu_1(R)k(R) \\
&+ \int_R^T [\mathcal{H}_2(s, R, c^*(s), k^*(s), \mu_2(s)) - \dot{\mu}_2 k(s)]ds + \mu_2(R)k(R) - \mu_2(T)k(T) \\
&+ \frac{\eta[T-R]^{1-\frac{1}{\phi}}}{1-\frac{1}{\phi}}.
\end{aligned}
\tag{29}
$$

Since $k(0) = k(T) = 0$ by assumption, and $\mu_1(R) = \mu_2(R)$ following the transversality condition, Equation (29) simplifies to:

$$
\begin{aligned}
V = &\int_0^R [\mathcal{H}_1(t, c^*(t), l^*(t), k^*(t), \mu_1(t)) - \dot{\mu}_1 k(t)]dt \\
&+ \int_R^T [\mathcal{H}_2(s, R, c^*(s), k^*(s), \mu_2(s)) - \dot{\mu}_2 k(s)]ds + \frac{\eta[T-R]^{1-\frac{1}{\phi}}}{1-\frac{1}{\phi}}.
\end{aligned}
\tag{30}
$$

The first order condition characterizing the decision for optimal retirement age then becomes:

$$
\frac{\partial V}{\partial R} = \mathcal{H}_1(R) - \mathcal{H}_2(R,R) + \int_R^T \frac{\partial \mathcal{H}_2(s,R)}{\partial R} ds - \eta[T-R]^{-\frac{1}{\phi}} = 0. \tag{31}
$$

Equation (31) is in principle the standard condition for the optimal switching point in a two-stage control problem (see e.g., Kamien and Schwartz, 2012), augmented with the marginal effect of a small change in R on the retirement income and the retirement good. Substituting the specifications of the Hamiltonian functions into Equation (31), the explicit condition for R^* becomes:

$$
\begin{aligned}
&\beta ln(l^*(R))e^{-\theta R^*} - \eta[T-R^*]^{-\frac{1}{\phi}} + \mu_0 \Bigg\{ (1 - l^*(R^*))w(1-\tau)e^{-rR^*} - b(R^*)e^{-rR^*} \\
&\quad + \frac{\tau}{(T-R^*)^2} \Bigg[\kappa w_i \Bigg((T-R^*)(1-l^*(R^*)) + \int_0^{R^*} (1 - l^*(t))dt \Bigg) \\
&\quad + (1-\kappa)Y \Bigg] \int_R^T e^{-rs} ds \Bigg\} = 0.
\end{aligned}
\tag{32}
$$

Rearranging the terms in Equation (32), one can achieve the condition for optimal retirement age as expressed in Equation (7).

Appendix B. Numerical solver

Since the optimization problem of both individual types are interdependent via the pension benefit formula, I solve the model by employing the following iterative process:

1. Guess the value of the aggregate output, Y_{guess};
2. Given Y_{guess}, solve the optimal control problem of the individuals;
3. Given obtained values for $l_i^*(t)$, and R_i^* from step 2, calculate a new value for aggregate output, $Y_{feedback}$;
4. Replace Y_{guess} with $Y_{feedback}$ and iterate on Y_{guess} until $(Y_{feedback} - Y_{guess})^2 < 0.000001$.

When the iteration process has converged according to the criterion in step 4, I consider the model to be in equilibrium where no individual would benefit from making any adjustments on their margins of decision.